**Noname manuscript No.**
(will be inserted by the editor)

---

# Enhancing spatial ability and mechanical reasoning through a STEM course

**First Author · Second Author**

Received: date / Accepted: date

**Abstract** There is a clear contemporary interest for developing science, technology, engineering, and mathematics (STEM) at schools. Besides, there exist a lot of research that justify the importance of spatial ability to obtain success in STEM subjects. Nevertheless, the spatial ability is relatively ignored in the general practice of teaching and learning in the K-12 setting. The goal of this paper is to analyse the evolution of spatial abilities of students that assist to a STEM course. Additionally, the evolution of their mechanical reasoning is also analysed. The STEM course was designed and implemented for the first time in a 6th grade class (primary school) and a 7th grade class (secondary school) throughout a whole academic year. First, the spatial ability and the mechanical reasoning of the students were assessed with the corresponding prepared pre-tests. Then, after finishing the STEM course, the students were tested with analogous post-tests. An exhaustive analysis of the obtained results is provided in the paper. It is shown that the spatial ability of the students was definitely improved. Furthermore, this improvement was statistically significant. Results also evidenced that the mechanical reasoning of the participants was also improved, although the improvement was not statistically significant. Moreover, this research showed that, in general, obtained results do not depend significantly on the gender of the participants. Finally, results manifested the statistically significant difference of spatial ability between 6th grade and 7th grade students. The difference between grade levels was not as significant in the mechanical reasoning case.

---

F. Author
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

S. Author
second address

**Keywords** Spatial ability · Mechanical reasoning · STEM · Primary and Secondary education

## 1 Introduction

There is not a clear consensus to define the spatial ability concept. Indeed, several researchers have asserted that spatial ability is not a single skill, but it comprises multiple factors. Lin [17] summarizes several proposals existing in the literature that characterize different factors or categories of that concept. On the one hand, McGee [24] considers that spatial ability comprises at least two components: spatial visualization and spatial orientation. Lohman [20], on the other hand, divides spatial ability into three factors, namely spatial visualization, spatial orientation and spatial relations. A different definition is proposed by Linn and Petersen [18], which describe three categories of spatial ability labeled as spatial perception, mental rotation and spatial visualization. In [10], Carrol proposes that spatial ability comprises five factors: spatial visualization, spatial relations, visuospatial perceptual speed, closure speed and closure flexibility. In a more recent approach, Sutton and Williams [30] define spatial ability as the performance on tasks that require mental rotation of objects, the ability to describe and understand how objects appear at different angles and understanding of how objects relate to each other in space.

The National Council of Teachers of Mathematics (NCTM) [1] set goals for mathematics education in the prekindergarten through to grade 12. Focussing on the *Geometry Standard*, they point out that geometry has long been regarded as the place in high school where students learn to prove geometric theorems[1]. The *Geometry Standard* they propose takes a broader view of the power of geometry by calling on students to analyse characteristics of geometric shapes and construct mathematical arguments about the geometric relationships, as well as to use visualization, spatial reasoning, and geometric modeling to solve problems.

In addition to the NCTM's recommendation to introduce visualization and spatial abilities at the primary school level, several authors insist on the importance of acquiring these abilities in order to have academic success in areas such as science, technology or engineering (e.g., [29–31, 11]).

Indeed, the identification and development of STEM talent has become a national priority in most developed countries (e.g, [8]). Actually, the supply of highly qualified scientists, technologists, engineers and mathematicians is perceived by governments globally as being vital in securing economic prosperity [5].

Ritz and Fan [28] present an international state-of-the-art of the STEM and technology education. Particularly, they report the perceptions of 20 international technology education scholars on their country's involvement in science, technology, engineering, and mathematics (STEM) education. They

---

[1] http://www.corestandards.org/Math/

comment that the implementation of STEM has resulted in varied approaches with each having potential difficulties for long-term success. Specifically, many discussions of STEM have focused on the improved teaching of separate subjects, especially those of science and mathematics. In some cases, STEM has incorporate technology and engineering into its framework as a means to show how scientific applications can be incorporated within science and mathematics lessons. In other cases, STEM is thought as a multidisciplinary approach to learning that can be used to integrate knowledge from the separate STEM subjects into existing science, technology, engineering design-based studies, or problem-based learning strategies. Some see it taught as a new integrative subject labeled as STEM. In this interpretation, educational practices would use multiple STEM subjects and integrate these into single courses.

According to [28], some of the reasons that the concept of STEM education has emerged are the lack of students progress in STEM school subjects and their election not to pursue these professions as careers, and economic goals for countries projected by politicians. However, several reports have pointed out current failures of educational systems in helping students to understand how to solve real-world problems using knowledge gained through the study of science, technology, engineering, and mathematics studies( [9], [2]). It appears many students are losing their potential competitiveness for the high-tech knowledge-based economies, because of their low performances in and dislikes of these subjects. As a result, STEM educational reform has become a topic of discussion in political, economic, and educational circles around the world.

Considering the contemporary push for developing STEM (science, technology, engineering, and mathematics), Lubinski [22] proposes to incorporate spatial ability in talent identification procedures. The findings the author exposes suggest that individuals who go onto achieve educational and occupational credentials in STEM tend to be distinguished by salient levels of spatial ability. Moreover, the author affirms that spatial ability covaries with preference patterns correspondent with the motivational (interest-values) profiles of individuals with STEM degrees and occupations.

Andersen [3] also remarks that the visual-spatial ability has predictive validity for future achievement in science, technology, engineering and mathematics (STEM) occupations. The author proposes to include visual-spatial ability in gifted education. Furthermore, Andersen claims that modern scientists need visual-spatial abilities to reconcile visual displays of real-world data with scientific models: data and models are transformed in their minds to find possible solutions. Thus, visual-spatial abilities are important for the STEM disciplines because many problems are solved through the creation of novel visualizations or mental model manipulation.

Notwithstanding all the existing research that defends the importance of the spatial ability to obtain success in STEM subjects, it is a capability rarely measured and is relatively neglected in the general practice of teaching and learning in the K-12 setting, as pointed out in [3].

In addition to enforce the spatial ability, hands-on learning proposals that promote mechanical reasoning or analogical reasoning also enable to foster

STEM disciplines. For instance, McKenna and Agogino [25] present the Simple Machines Learning Environment (SIMALE) to support mechanical reasoning and understanding of simple machines for middle school and high school students and show its effectiveness in student learning. Their primary goals were to foster development of simple machines concepts, and to encourage students to make connections between abstract and more concrete forms of reasoning in order to effectively apply their knowledge to a range of problems.

Liu and Schunn [19] show that higher levels of mechanistic knowledge are associated with more frequent and complex mathematical strategy use. Hence, the authors suggest that mechanistic knowledge may be on pathway through which adaptive mathematical strategy use can be improved. The participants are asked to program a robot to navigate a maze and to create a navigation strategy that would work for differently sized robots.

Similarly, Cuperman and Verner[12] propose to foster analogical reasoning and design skills through creating bio-inspired robotic models. The authors found, among others, that the students acquire technological content knowledge essential or using the construction kit and constructing simple robotic systems. Furthermore, they are highly motivated to learn scientific and technological concepts and perform hands on activities.

Taking into account the previously exposed ideas, the current research aims at promoting the spatial abilities and mechanical reasoning through a course labeled as STEM. Remark that each of the aforementioned proposals considers different factors of spatial ability. However, most of them include spatial visualization, spatial orientation or spatial relations. This study considers these three factors to define spatial ability, which are precisely the proposed ones by Lohman [20].

## 2 The research study

In a pilot project published in [7], we analysed the use of educational robotics to develop spatial abilities in 12 year old students. To carry out that project, a course to introduce robotics to 6th grade primary school students was designed. With only an 8 sessions-course, obtained results showed that the students that joined the course improved significantly their spatial abilities compared to the students that did not join the course.

Additionally, the teacher who implemented that project noticed the high levels of motivation and interest of the students for the robotics course. He also noticed (although he did not demonstrate it) that the students did not only learn concepts related to mathematics, but also related to science, engineering and technology.

Based on the good results obtained in that pilot project, the school decided to design a more general course. Particularly, an integrative subject, that contains science, technology, engineering and mathematics, and that is labeled as STEM course was designed. In the aforementioned pilot project, the school only had 3 kits of material and only 9 students were able to join the robotics course. In order to implement this new integrative subject, the school bought more kits of material and more varied ones. Hence, in the STEM course not only the robotics could be studied but also technology and engineering. The course was compulsory for all the students and was taught during a whole academic year. In this first year of implementation, the STEM course was included in the annual planning of two different grades: 6th grade (primary education) and 7th grade (secondary education). As mentioned in [8], engaging elementary and middle school students has the greatest impact on closing the STEM educational gap.

### 2.1 Objective

The main objective of this paper is to analyse if the students that attended the STEM course improve their spatial ability. Furthermore, this work also aims at studying if the students enhance their mechanical reasoning.

Specifically, the current research was designed to answer the following questions:

- Do the students that attend the STEM course improve their spatial ability?
- Do the participants in the STEM course enhance their mechanical reasoning?
- Do these results depend on the gender of the participants?
- Do these results depend on the grade level?

The remainder of the paper is organized as follows. First, the material used in the course is briefly presented. Then, the sample and methodology are introduced. Next, the instruments that allows to quantify the acquisition of spatial ability and of mechanical reasoning are detailed. Furthermore, examples of hands-on activities proposed during the sessions are included. Section 3 provides the obtained results and a discussion considering each of the research questions. Finally, conclusions and future work are given in Section 4.

### 2.2 Material

The material used in the sessions consisted of 7 different Fischertechnik[2] sets. Table 1 summarizes the distribution of the sets in each of the grades (see Fig. 1 for illustrations of sample sets). Each kit, which allows to construct several models, aims at fostering different skills. An example of a model corresponding to each of the kits is shown in Fig. 2.

The ROBOTICS sets include the ROBO pro Light software, which is used to program the model robots. A snapshot of the software is shown in Fig. 3. The Oeco Energy set aims at showing how the electricity can be produced, storages and used from natural energy sources such as water, wind and the sun. This set includes a solar motor (2V) and 2 solar modules (1V; 400 mA).

---

[2] http://www.fischertechnik.de/en/home.aspx

**Table 1** Sets used in the STEM course

| 6th | | 7th | |
|---|---|---|---|
| ROBOTICS LT Beginner Set | 4 | ROBOTICS TXT Discovery Set | 4 |
| Oeco Energy | 2 | Oeco Energy | 2 |
| Mechanic + Static | 2 | Pneumatic 3 | 2 |
| Optics | 2 | Dynamic XL | 2 |

**Fig. 1** Material used in the 6th grade (top row) and in the 7th grade (bottom row)

**Fig. 2** Examples of the models constructed with the material used in the 6th grade (top row) and in the 7th grade (bottom row)

The Mechanic and Static set is suitable for all future engineers and technicians. The provided models show how a shaft drive or manual transmission works or how a stable bridge can be constructed, etc. The Optics set is provided with optical lenses with various focal lengths, mirrors, lens tip lamps and a variety of other parts that allow construction of a microscope, magnifier, telescope and periscope. The Pneumatic set takes a learning-by-playing approach to the principles of pneumatics, with realistic models to show how pneumatic valves

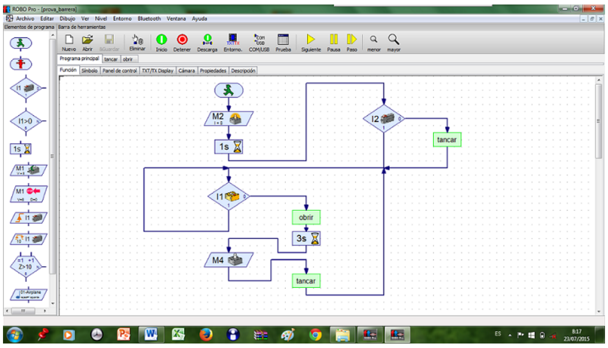


**Fig. 3** Snapshot of the ROBO Pro Light 4.0 software

and cylinders work. Finally, the Dynamic XL allows to play with a ball that races through tight curves and chutes and shoot through different tracks.

## 2.3 Sample

The study described in this paper was conducted in a school in a small city in Spain. The study was carried out with a 6th grade class and a 7th grade class. There were 26 students in the 6th grade class (15 boys and 11 girls) and 24 students in the 7th grade class (11 boys and 13 girls). The students worked in 3-members teams. Each team had a computer to work with, in case the particular kit requires it. As in [12], a team of three learners was found preferable to enable self-expression and active participation, while still allowing the benefits of team collaboration.

As mentioned above, the STEM course was compulsory for all the students and was taught in a 1-h/week setting throughout all the 2015-16 academic year (Monday, 4:00-5:00pm for 7th grade and Thursday, 4:00-5:00pm for 6th grade). A total of 26 and 28 sessions were developed in the 7th and 6th grade, respectively.

An important consideration of the STEM course was that the classroom environment should be a large space. Therefore, the science laboratory was used (see Fig. 4). There were big rectangular tables that were suitable for working in groups. Furthermore, the space was very well illuminated, which is crucial when working with such a large amount of different components. Additionally, the stools allow a better position to work.

Table 2 summarizes the main differences of the current research settings with respect to the ones of the pilot project [7].

## 2.4 Methodology

The main purpose of this study was to determine if the STEM course helps to improve the spatial ability of the participants. Furthermore, another goal of

**Fig. 4** Classroom where the STEM course was implemented

**Table 2** Main differences between the pilot project and the current research

| | pilot project | current research | |
|---|---|---|---|
| | 6th | 6th | 7th |
| participants | 9 | 26 | 24 |
| sessions | 8 | 28 | 26 |
| kits | 3 | 4 | 4 |
| research questions | spatial ability evolution | spatial ability evolution, mechanical reasoning evolution, gender differences, grade level differences | |

the current research was to analyse if the students enhance their mechanical reasoning. These two evaluations can be performed by means of a test procedure. Details on the characteristics of the tests are provided in the following section. Therefore, at the beginning of the first term (September 2015) the students completed a test (pre-test) to determine their spatial ability and a test (pre-test) to evaluate their mechanical reasoning. Then, at the end of the third term (June 2016), the students completed a test (post-test) to detect if the spatial ability has been improved significantly and a test (post-test) to study if their mechanical reasoning has been enhanced. Moreover, this research analyses if the results depend on the gender of the participants. Finally, the relation between obtained results and the students' grade (6th or 7th) is also studied.

As described above, the students worked in 3-members teams. The idea was that at each session they had to solve a particular problem. In general, they had to perform a model and to answer some questions related to it. The teacher's role was as a guide and the students had to face their activities working in group and through a hands-on learning.

As in [25] and [19], the idea was to provide the students with real problems and machines in order to motivate them to use mathematical tools they already know to solve real life situations. In addition, the characteristics that a learning environment should have according to [25] were taken into account. They proposed that it should provide opportunities for students to actively participate, support self-reflection, provide multiple representations of concepts, and cultivate generative learning.

The sessions were planned in order to allow the rotation of the sets among the working teams. That is, at the end of the STEM course, all the teams will have worked with every kit.

It should be highlighted that the co-author of this paper was the designer of the sessions and the teacher in the STEM course presented in this research. The co-author of this paper is a primary school teacher and also an industrial engineer. His multidisciplinary educational background enables him to easily teach technology and engineering at primary school. In addition, he had experience in teaching the robotics course previously implemented in the pilot project.

## 2.5 Instruments

### *2.5.1 Spatial ability*

There exists a wide range of literature that aims at evaluating spatial ability through the use of tests. Some of the most popular tests according to [11, 27, 29] are the following: Mental Rotation Test (MRT); Differential Aptitude Test-Spatial Relations (DAT-SR); Mental Cutting Test (MCT); Purdue Spatial Visualization Test (PSVT:R); and Vandenberg MRT. The problem is that all these tests are not designed for 12 year old students, but for older ones. Therefore, they were deemed not suitable for the current study.

In addition, since spatial ability comprises multiple factors, it could be better to use several sub-tests instead of using a single test, as proposed by Humphreys [16]. As in the pilot project [7], the instrument used in this study to evaluate the spatial ability of the students (pre- and post-test) consisted of 4 sub-tests. Three of the considered sub-tests were based on Bakker's proposal ([4]), in which the author analyses different tests for evaluating the spatial ability of 11 year old students. The other sub-test considered for use in the current study was selected from the Spatial Intelligence and Learning Center[3], which is a web site that provides sets of tests to evaluate spatial skills. Specifically, the selected sub-test was the *Perspective Taking/Spatial Orientation Test* ([15]).

Table 3 summarises the number of items of each sub-test. Note that their items are split into two parts, corresponding to the pre- and post-test, respectively. The students have 3 minutes to finish sub-tests 1, 2 and 4, which have 10, 10 and 21 items, respectively, and 2 min and 30 s to finish the sub-test 4, which has only 6 items.

These four sub-tests selected in this paper aims at evaluating the different factors of spatial ability considered in the current research: spatial visualization, spatial orientation, and spatial relations. Specifically, the *Card Rotations Test* and the *Cube Comparison Test* allow to evaluate spatial relations, as considered in [14] and [10]. The *Perspective Taking Spatial Orientation Test*

---

[3] http://www.spatiallearning.org/

**Table 3** Pre- and post-test composition

| sub-test | name | parts | items |
|---|---|---|---|
| 1 | *Card Rotations Test* ([13]) | 2 | 20 |
| 2 | *Paper Folding Test* ([13]) | 2 | 20 |
| 3 | *Perspective Taking Spatial Orientation Test* ([15]) | 1 | 12 |
| 4 | *Cube Comparisons Test* ([13]) | 2 | 42 |

**Fig. 5** An item of the *Card Rotations Test*

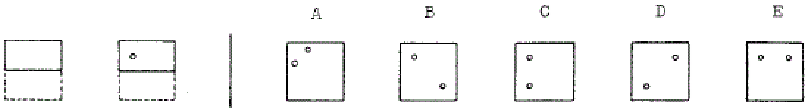


**Fig. 6** An item of the *Paper Folding Test*

enables to evaluate the spatial orientation. Finally, the *Paper Folding Test* allows to evaluate spatial visualization, as proposed in [21] and [18].

A brief explanation of each sub-test and a relevant example of its contents is provided below.

*Card Rotation Test.* This test requires mental rotations of objects. The students have to decide if the objects on the right correspond to the object on the left, in which case they mark S (same). They mark D (different), if otherwise (see the sample responses given in Fig. 5).

*Paper Folding Test.* The students have to imagine how to fold and unfold a sheet of paper. Instructions to fold the paper are given on the left (see Fig. 6). Then, a hole is made in the paper. Once unfolded, a single figure on the right corresponds to the original paper on the left (in the example given in Fig. 6, the correct answer is C).

*Perspective Taking/Spatial Orientation Test.* This test requires the visualization of different perspectives and orientations of objects in space. An example is shown in Fig. 7. There are always the same objects on the top. Information about the position of the student is provided in the middle. On the bottom, the student has to translate the given information to a scheme.

*Cube Comparison Test.* This test requires mental rotations of objects in 3D. The cubes contain a different symbol in each face. The students have to decide

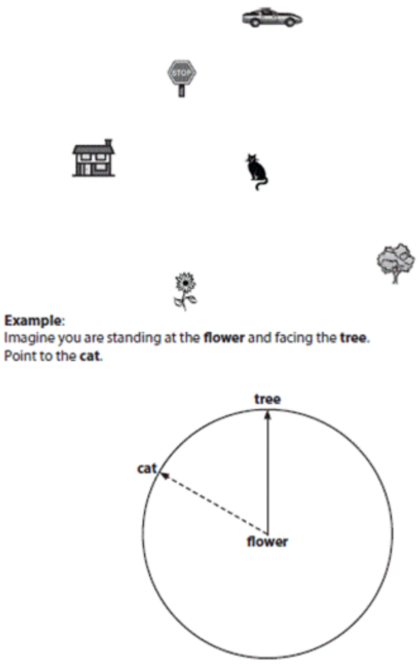

**Example:**
Imagine you are standing at the **flower** and facing the **tree**.
Point to the **cat**.

**Fig. 7** An item of the *Perspective Taking/Spatial Orientation Test*. The instructions given in the middle are: *You are standing at the flower facing the tree. Point to the cat.* The dashed line on the bottom corresponds to the expected answer

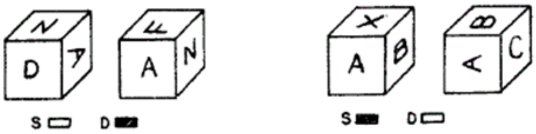


**Fig. 8** Two items of the *Cube Comparisons Test*

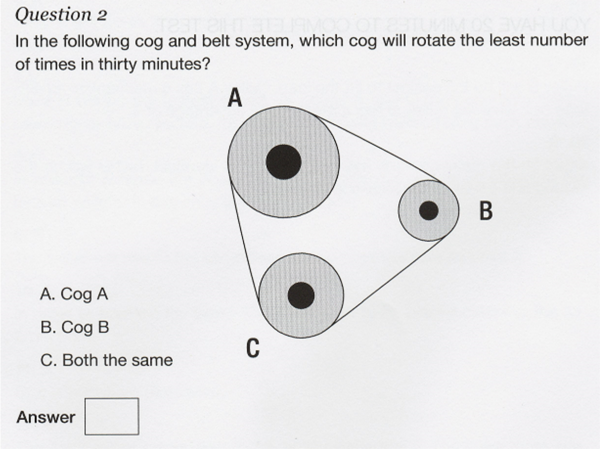
if the two given images correspond to the same cube (see two examples in Fig. 8).

The steps to answer each of the sub-tests were the following:

- The teacher stated the name of the sub-test and explained it briefly.
- The teacher showed an example of the problem using a projector.
- The teacher answered doubts.
- The students completed the sub-test.

### *2.5.2 Mechanical Reasoning*

The mechanical reasoning can be measured through the Bennett Mechanical Comprehension Test (BMCT). As pointed out in [6], the BMCT is an



*Question 2*

In the following cog and belt system, which cog will rotate the least number of times in thirty minutes?

A. Cog A

B. Cog B

C. Both the same

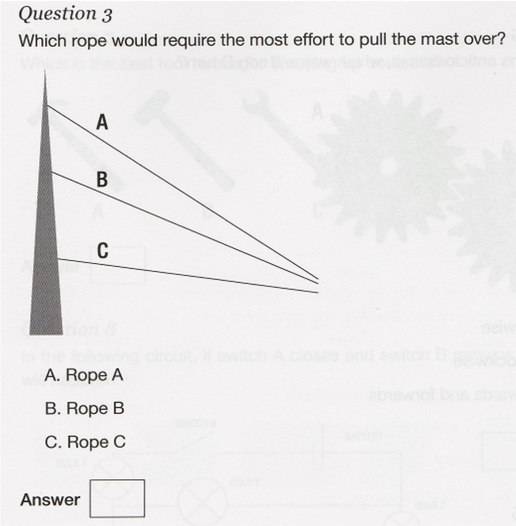
Answer ☐

**Fig. 9** An item of the *Mechanical Reasoning Test*

aptitude test used to measure a person's ability for future learning. Particularly, this test was developed to measure a person's aptitude for understanding and applying mechanical principles, from which an employer may infer future performance in jobs that require these skills. This aptitude, known as mechanical comprehension, is regarded as one aspect of intelligence, as intelligence is broadly defined. The individual who scores high in mechanical comprehension tends to learn readily the principles of the operation and repair of complex devices. Like those who take other aptitude tests, a person's performance on the BMCT may be influenced by environmental factors, but not to the extent that interpreting his or her performance is significantly affected. Although an individual's scores on the BMCT can generally be improved through training and experience, it is unlikely that improvement will be dramatic. This situation is due in part to the presentation and composition of items that are simple, frequently encountered mechanisms, neither resembling textbook illustrations nor requiring special knowledge.

The *Mechanical Reasoning Test* used in the current research are resulted from a selection of questions included in [26], which aims at preparing students for the BMCT. Fig. 9 and 10 show two items of the *Mechanical Reasoning Test*. The correct answers are A and C, respectively.

It is worthy to remark that none of the activities included in the tests presented above were directly addressed in the STEM course. Therefore, the students were not trained to answer the items of the spatial ability tests nor the ones of the mechanical reasoning test. These instruments were selected to measure if the STEM course implicitly helps to improve the spatial ability and the mechanical reasoning of the students.



*Question 3*

Which rope would require the most effort to pull the mast over?

A. Rope A

B. Rope B

C. Rope C

Answer ☐

**Fig. 10** An item of the *Mechanical Reasoning Test*

## 2.6 Hands-on learning proposals

This section provides two examples of problems that the students should solve. The goal at this stage is to illustrate the hands-on learning essence of the problems and activities proposed during the sessions of the STEM course.

Fig. 11 shows students working in groups. Each group was assigned a particular kit together with the Fischertechnik's original manual. Furthermore, the teacher provided documentation to each of the teams containing the information about the main characteristics of the assigned kit, the components it contains and some useful instructions to understand how they work. The documentation also included a problem the team had to solve during the session. Each of the team had a different problem to solve during the session. This fact often generated discussions among the teams, which used to be interested in the models their classmates were preparing.

An example of a problem given to a group that works with the ROBOTICS TXT Discovery set is provided in the box below.

**Fig. 11** Students working in 3-members groups

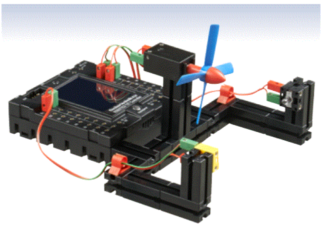


**Fig. 12** Hand dryer model, ROBOTICS TXT Discovery set

***Task III:*** *Build the hand dryer model shown in Fig. 12 according to the instructions of the Fischertechnik's manual (page 10).*

- *The hand dryer should be programmed according the following instructions: when the light barrier is interrupted, the fan must be connected. Then, after 10 seconds, it must be automatically disconnected (File name: Task3a.rpp).*
- *Modify the program of the previous section. In this new version, during the first 5 seconds the fan should rotate at his maximum velocity and during the last 5 seconds it should rotate at a moderate velocity (File name: Task3b.rpp).*
- *Since the goal is to save energy, create a new program in which the fan turns on when hands approach the dryer and it turns off when hands go away from it (File name: Task3c.rpp).*

An example of a problem given to a team that works with the Mechanic and Static kit is provided below.

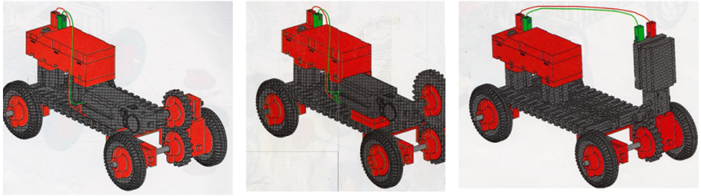


**Fig. 13** Different vehicle models, Mechanic and Static set: (left) model 1; (middle) model 2; (right) model 3

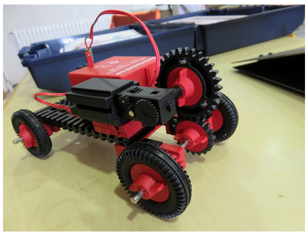


**Fig. 14** Vehicle model 2 built by the students

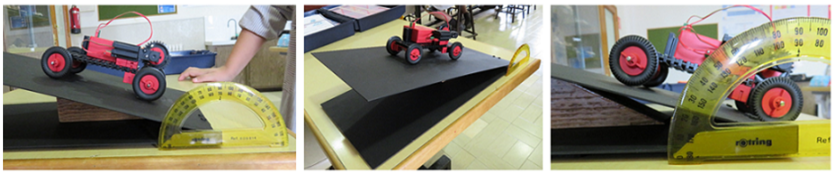


**Fig. 15** Vehicle model 2 going up the ramp

***Task V:***

- *Build the vehicle model 1 shown in Fig. 13 (left) according to the instructions of the Fischertechnik's manual (page 14).*
- *Build the vehicle model 2 shown in Fig. 13 (middle) according to the instructions of the Fischertechnik's manual (page 17).*
- *Build the vehicle model 3 shown in Fig. 13 (right) according to the instructions of the Fischertechnik's manual (page 19).*
- *For each of the vehicle models, detail:*
  ***Q1:*** *its transmission relation,*
  ***Q2:*** *the time it spends travelling 5 meters,*
  ***Q3:*** *if it is able to go up a ramp.*

Fig. 14 shows the model 2 built by the corresponding team. This model was able to go up a ramp of 21°, as it is shown in Fig. 15. Table 4 summarizes the solution provided by the corresponding team. They deduced that the vehicle model 2 is the fastest one. They also concluded that they needed to adapt the ramp to the characteristics of each model.

**Table 4** Solution given by the team working with the task V

| question | model 1 | model 2 | model 3 |
|---|---|---|---|
| Q1 | 1 | 1.5 | 0.5 |
| Q2 | 6.19 sec | 4.28 sec | 11.44 sec |
| Q3 | yes, 24° | yes, 21° | yes, 28° |

It is important to remark that the students arrived to a solution working as a team. It is also significant to notice that in order to be able to solve the assigned tasks or problems, they must understand concepts and procedures related to engineering and technology (e.g., when using the ROBO Pro Light software, or when calculating the transmission relation of each of the models).

## 3 Results

This section provides an in-depth analysis of the obtained results in order to answer the research questions formulated above. The evolution of spatial ability and mechanical reasoning are studied separately. In each of the cases, the number of participants is different, since only the students who took the corresponding pre- and post-tests were considered for the analysis. Specifically, the number of students that completed both spatial ability tests is 25 and 24 in the 6th and 7th grades, respectively. In the mechanical reasoning case, 26 6th grade students and 22 7th grade students took both tests.

### 3.1 Spatial ability

#### *3.1.1 Evolution*

Fig. 16 shows the individual students average scores in the pre- and post-tests in the case of 6th grade (notice that the mean of the scores obtained in the 4 sub-tests gives the average score). Concretely, the score corresponds to the percentage of correct answers. The unanswered questions were computed as wrong. In particular, the polygonal figures in Fig. 16 enclose data in between lower and upper quartiles (medians are represented by horizontal lines in thinner regions). Additionally, the scores corresponding to female and male participants are shown. It can be seen that the scores obtained in the post-test are clearly higher than the ones obtained in the pre-test, in all the cases.

In order to compare the global performance of the 6th grade students in the pre- and post-tests, Table 5 summarizes the mean and standard deviation scores obtained by the whole 6th group in both tests. Specifically, results obtained in each of the 4 sub-tests and on average are detailed. It can be seen that the mean score obtained in each of the 4 sub-tests increases from the pre- to the post-test.

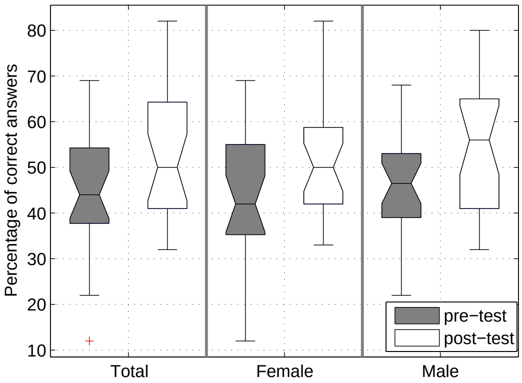


**Fig. 16** Individual students average scores in the tests, 6th grade students

**Table 5** Pre- and post-test scores (mean and standard deviation of the whole 6th grade group)

| | Mean (%) | | SD (%) | | | |
|---|---|---|---|---|---|---|
| Sub-test | Pre | Post | Pre | Post | t | p |
| 1 | 58.60 | 66.16 | 22.72 | 23.38 | -1.16 | 0.2521 |
| 2 | 40.40 | 47.20 | 23.35 | 19.04 | -1.13 | 0.2649 |
| 3 | 21.32 | 37.96 | 20.09 | 32.80 | -2.16 | **0.0356** |
| 4 | 56.32 | 58.88 | 12.21 | 9.09 | -0.84 | 0.4048 |
| Average | 44.12 | 53.64 | 14.44 | 14.06 | -2.35 | **0.0225** |

Additionally, a t-test was performed to compare the means obtained in the pre- and post-tests. Particularly, a two-tailed test is performed, setting $\alpha = 0.05$ as a significance level (5%). It can be seen that the improvement in the performance of the students is different in each of the sub-tests. Notice that in the *Perspective Taking/Spatial Orientation test* (sub-test 3) there is a statistically significant gain on the post-test scores ($p = 0.0356$). Although in the other sub-tests $p > 0.05$, on average, the improvement of the scores obtained in the post-test is statistically significant ($p = 0.0225$).

Analogously, Fig. 17 shows the individual students average scores in the pre- and post-tests in the case of 7th grade. As in the 6th grade case, the average scores obtained in the post-test are definitely higher than the ones obtained in the pre-test, considering all the participants and also in the cases of female and male participants.

Table 6 summarizes the mean and standard deviation obtained by the whole 7th grade group in the pre- and post-tests, for the 4 sub-tests and also on average (last row of the table). As before, the improvement of the score obtained in the sub-test 3 (*Perspective Taking/Spatial Orientation test*) is statistically significant ($p = 0.0471$). In the sub-test 2 (*Card Rotations test*), the gain is near to be statistically significant ($p = 0.0586$). Although the improve-

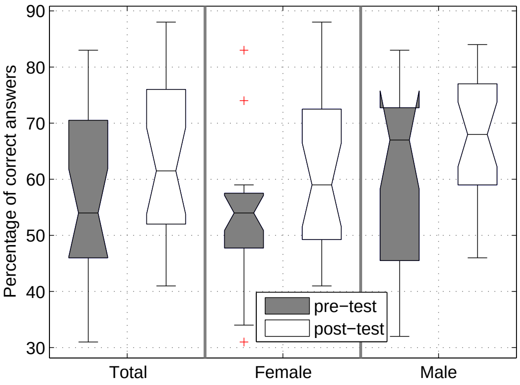


**Fig. 17** Individual students average scores in the tests, 7th grade students

**Table 6** Pre- and post-test scores (mean and standard deviation of the whole 7th grade group)

| | Mean (%) | | SD (%) | | | |
|---|---|---|---|---|---|---|
| Sub-test | Pre | Post | Pre | Post | t | p |
| 1 | 76.62 | 84.75 | 16.45 | 12.26 | -1.94 | 0.0586 |
| 2 | 48.33 | 55.41 | 21.19 | 18.64 | -1.23 | 0.2252 |
| 3 | 38.16 | 56.91 | 34.45 | 28.96 | -2.04 | **0.0471** |
| 4 | 57.95 | 59.33 | 13.61 | 10.28 | -0.39 | 0.6948 |
| Average | 56.25 | 63.95 | 15.54 | 13.75 | -1.82 | 0.0753 |

ment is not as significant in the other sub-tests, on average, the improvement of the scores obtained in the pre- and post-tests is near to be statistically significant ($p = 0.0753$).

### *3.1.2 Gender difference*

Fig. 18 shows that the average scores obtained in the female and male participants of 6th grade are similar in both the pre- and post-tests. However, the median obtained by the male participants is higher than the one obtained by the female participants.

Table 7 details the mean and standard deviation obtained by the female and male participants of the 6th grade group in the four sub-tests and also on average. In can be seen that, on average, the mean obtained by the male participants is higher than the one obtained by the female participants. There are only two exceptions in which the female participants obtain higher scores: pre-test of sub-test 2 and post-test of sub-test 3.

Additionally, a t-test was performed to compare the respective means obtained in the pre- and post-tests. Obtained $p$-values allow to conclude that the difference between the mean corresponding to the female and male students is

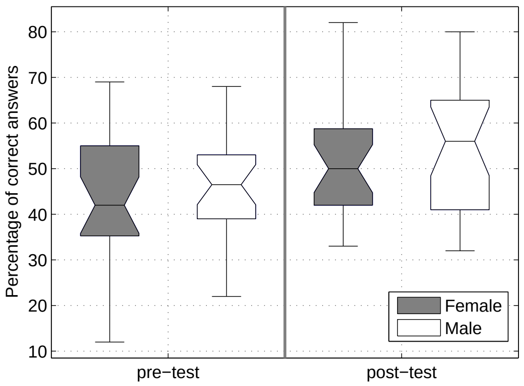


**Fig. 18** Individual students average scores in the tests, 6th grade students

**Table 7** Pre- and post-test scores (mean and standard deviation of the 6th group, considering the students gender)

| | Pre-test | | | | Post-test | | | |
|---|---|---|---|---|---|---|---|---|
| | Female (N=11) | | Male (N=14) | | Female (N=11) | | Male (N=14) | |
| test | M (%) | SD (%) | M (%) | SD (%) | M (%) | SD (%) | M (%) | SD |
| 1 | 53.36 | 27.22 | 62.71 | 26.78 | **55.63** | 18.49 | **74.42** | 17.03 |
| 2 | 44.54 | 27.33 | 37.14 | 22.96 | 45.55 | 20.16 | 48.57 | 16.10 |
| 3 | 19.72 | 19.40 | 22.57 | 32.83 | 39.36 | 21.26 | 36.85 | 33.98 |
| 4 | 53.63 | 8.29 | 58.43 | 9.61 | 58.45 | 14.54 | 59.21 | 9.01 |
| Average | 42.77 | 16.52 | 45.14 | 15.06 | 51.98 | 13.12 | 54.93 | 13.67 |

only significant in the post-test of sub-test 1, which is marked in bold in the table ($p = 0.0434$). In all the other cases, $p > 0.05$.

In the 7th grade case, the scores obtained by the male students are certainly higher than the ones obtained by the female students, as shown in Fig. 19.

Table 8 shows that the mean scores corresponding the male participants are higher than the ones corresponding to female participants except in the sub-test 2 (*Paper Folding test*). After performing a $t$-test, it can be concluded that the difference between the mean corresponding to the female and male students is only significant in the pre-test of sub-test 1, which is marked in bold in the table ($p = 0.0026$). In all the other cases, $p > 0.05$. Therefore, the difference between results obtained by female and male participants is not statistically significant, in general.

### *3.1.3 Level differences*

Fig. 20 shows the average scores obtained in the pre- and post-tests for each grade. The 7th grade students clearly obtain higher scores than the 6th grade ones.

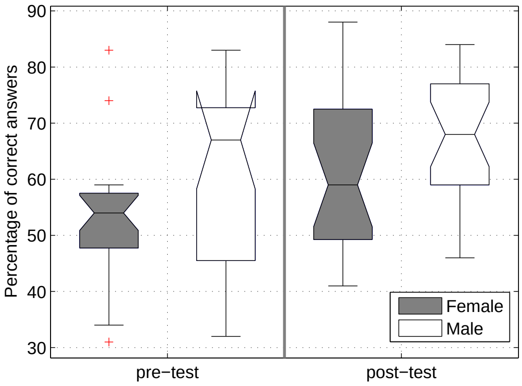


**Fig. 19** Individual students average scores in the tests, 7th grade students

**Table 8** Pre- and post-test scores (mean and standard deviation of the 7th group, considering the students gender)

| | Pre-test | | | | Post-test | | | |
|---|---|---|---|---|---|---|---|---|
| | Female (N=13) | | Male (N=11) | | Female (N=13) | | Male (N=11) | |
| test | M (%) | SD (%) | M (%) | SD (%) | M (%) | SD (%) | M (%) | SD |
| 1 | **67.92** | 16.37 | **86.91** | 13.96 | 80.61 | 9.27 | 89.63 | 7.97 |
| 2 | 50.77 | 19.77 | 45.45 | 18.43 | 56.92 | 23.39 | 53.63 | 19.63 |
| 3 | 29.54 | 28.12 | 48.36 | 31.77 | 50.00 | 39.62 | 65.09 | 24.14 |
| 4 | 57.61 | 13.13 | 58.36 | 7.90 | 58.61 | 14.80 | 60.18 | 12.91 |
| Average | 53.30 | 14.04 | 59.72 | 14.84 | 61.38 | 17.14 | 67.00 | 12.33 |

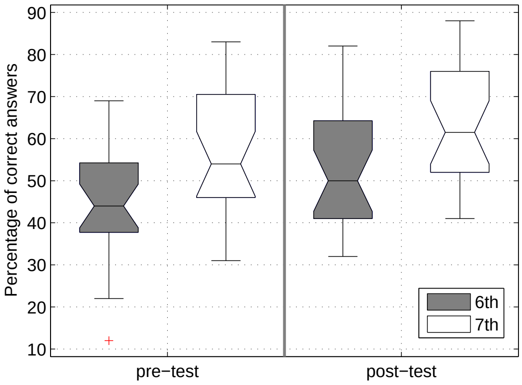


**Fig. 20** Individual students average scores in the tests

Table 9 summarizes the mean and standard deviation scores obtained in the post-test by the whole 6th and 7th groups (analogous results were obtained

**Table 9** Post-test scores (mean and standard deviation, 6th and 7th grade comparison)

| | 6th grade (N=25) | | 7th grade (N=24) | | | |
|---|---|---|---|---|---|---|
| Test | M (%) | SD (%) | M (%) | SD (%) | t | p |
| 1 | 66.16 | 23.38 | 84.75 | 12.26 | -3.4629 | **0.0011** |
| 2 | 47.20 | 19.04 | 55.41 | 19.04 | -1.5253 | 0.1339 |
| 3 | 37.96 | 32.80 | 56.91 | 28.96 | -2.1407 | **0.0375** |
| 4 | 58.88 | 9.09 | 59.33 | 10.28 | -0.1637 | 0.8707 |
| Average | 53.64 | 14.06 | 63.95 | 13.77 | -2.5936 | **0.0126** |

in the pre-test). Again, a *t*-test is performed to study if the difference between 6th and 7th grades is statistically significant. Obtained results allow to affirm that the difference between the means corresponding to the sub-test 1, the sub-test 2 and the average is statistically significant ($p = 0.0011$, $p = 0.0375$ and $p = 0.0126$, respectively). Average scores obtained in sub-test 2 and 4 are similar in both the 6th and 7th grade and $p > 0.05$.

## 3.2 Mechanical Reasoning

### *3.2.1 Evolution*

Fig. 21 shows the 6th grade students scores obtained in the mechanical reasoning test in both the pre- and post-tests. Scores obtained by female and male participants are also shown. The plot allows to confirm that post-test scores are in general higher than the pre-test scores. Notice the median value obtained in each case.

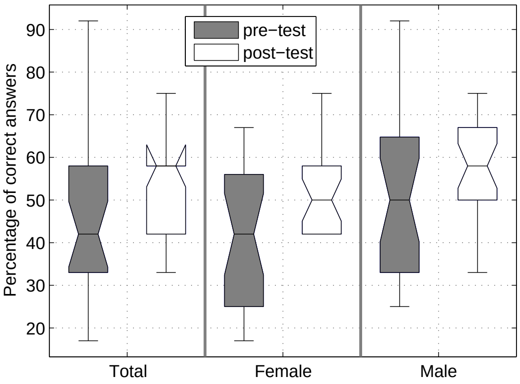


**Fig. 21** Individual students scores in the mechanical reasoning test, 6th grade group

**Table 10** Pre- and post-test scores (mean and standard deviation of the whole 6th grade group)

| | Mean (%) | | SD (%) | | t | p |
|---|---|---|---|---|---|---|
| | Pre | Post | Pre | Post | | |
| Total (N=26) | 47.15 | 54.50 | 19.85 | 11.07 | -1.6477 | 0.1057 |
| Female (N=11) | 41.09 | 50.81 | 18.40 | 10.62 | -1.5177 | 0.1447 |
| Male (N=15) | 51.60 | 57.20 | 20.29 | 10.94 | -0.9407 | 0.3549 |

Table 10 presents the mean and standard deviation of the scores obtained by the whole 6th grade group. It can be seen that the mean obtained in the post-test is higher than the one obtained in the pre-test, in all the cases. In addition, Table 10 includes the $p$-value computed by applying a $t$-test. Recall that, although $p > 0.05$ in all the cases, the difference between the obtained means are near to be statistically significant when all the participants or female participants are considered ($p = 0.1057$ and $p = 0.1447$, respectively).

Analogously, Fig. 22 shows the 7th grade students scores obtained in the pre- and post-tests. As before, the median obtained in the post-test is undoubtedly higher than the one obtained in the pre-test in all the cases.

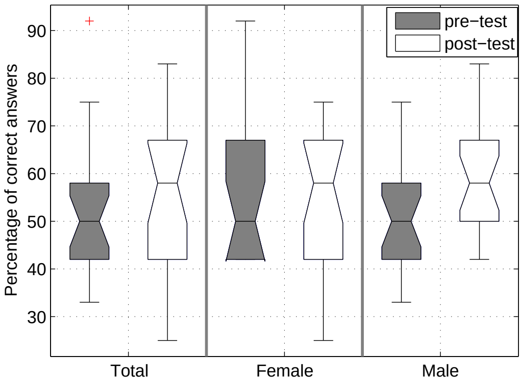


**Fig. 22** Individual students scores in the mechanical reasoning test, 7th grade group

Table 11 contains the mean and standard deviation of the scores obtained by the whole 7th grade group. Notice that the mean corresponding to the male participants increases in the post-test and also in the case all the participants are considered. The mean obtained by the female participants, on the contrary, decreases in the post-test (although the median value was clearly improved, as shown in Fig. 22). Furthermore, Table 11 also includes the results obtained after applying a $t$-test. Notice that the $p$-value is always higher than 0.05. However, in the male participants case, the difference between the

**Table 11** Pre- and post-test scores (mean and standard deviation of the whole 7th grade group)

| | Mean (%) | | SD (%) | | t | p |
|---|---|---|---|---|---|---|
| | Pre | Post | Pre | Post | | |
| Total (N=22) | 53.50 | 56.50 | 14.22 | 14.30 | -0.6974 | 0.4894 |
| Female (N=12) | 55.75 | 54.25 | 16.38 | 15.76 | 0.2285 | 0.8214 |
| Male (N=10) | 50.80 | 59.20 | 11.37 | 12.60 | -1.5649 | 0.1350 |

mean obtained in the pre- and post-tests is near to be statistically significant ($p = 0.1350$).

### *3.2.2 Gender difference*

Fig. 23 aims at comparing the female and male participants scores obtained in the pre- and post-tests in the 6th grade case. Scores obtained by the male participants are clearly higher that the ones obtained by the female participants, in both tests. Nevertheless, notice that the difference is not as notable in the post-test.

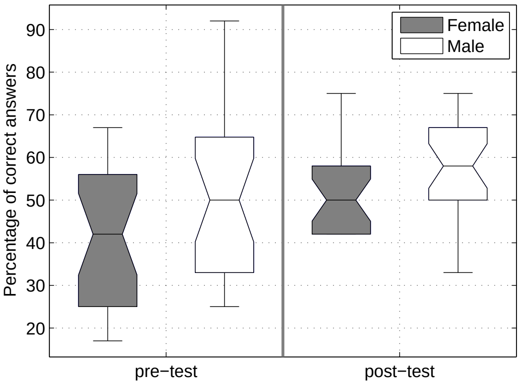


**Fig. 23** Individual students score in the test mechanical reasoning test, 6th grade group

Table 12 shows the mean and standard deviation of the scores obtained by the female and male participants of the 6th grade group. Additionally, it contains the results obtained after applying a $t$-test. It can be concluded that the difference between the mean obtained by the male participants and the obtained one by the female participants are near to be statistically significant ($p = 0.1879$ and 0.1501 in the pre- and post-tests, respectively).

Analogously for the 7th grade, Fig. 24 shows the female and male participants scores obtained in the pre- and post-tests. In this case, scores obtained

**Table 12** Gender difference in the pre- and post-test scores (mean and standard deviation, 6th grade group)

| | Female | | Male | | | |
|---|---|---|---|---|---|---|
| | Mean (%) | SD (%) | Mean (%) | SD (%) | t | p |
| pre-test | 41.09 | 18.40 | 51.60 | 20.29 | -1.3555 | 0.1879 |
| post-test | 50.81 | 10.62 | 57.20 | 10.94 | -1.4868 | 0.1501 |

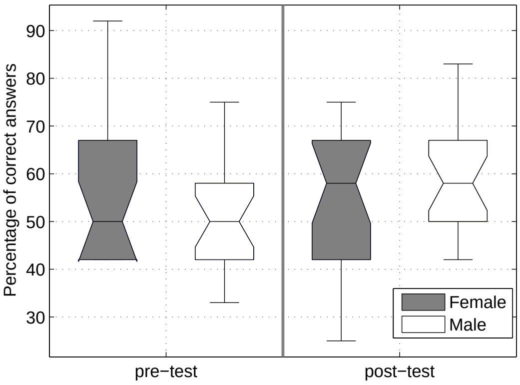


**Fig. 24** Individual students score in the test mechanical reasoning test, 7th grade group

**Table 13** Gender difference in the pre- and post-test scores (mean and standard deviation, 7th grade group)

| | Female | | Male | | | |
|---|---|---|---|---|---|---|
| | Mean (%) | SD (%) | Mean (%) | SD (%) | t | p |
| pre-test | 55.75 | 16.38 | 50.80 | 11.37 | 0.8057 | 0.4299 |
| post-test | 54.25 | 15.76 | 59.20 | 12.60 | -0.8012 | 0.4324 |

by the female and male participants are not as different as in the 6th grade case. Concretely, notice that in the pre-test, the female participants obtain higher scores than the male participants. In the post-test, on the contrary, the male participants obtain, in general, higher scores. Hence, it is clear that the evolution of the scores corresponding to female and male participants is different: the male participants improve their scores in a more significant way.

The mean and standard deviation obtained by female and male participants of the 7th grade in the pre- and post-tests are shown in Table 13. Notice that the difference between the mean obtained by the female and male participants is similar in the pre- and post-tests. However, in the pre-test, the mean is higher in the female case and, in the post-test, the mean is higher in the male case. Nevertheless, after applying a $t$-test, it can be concluded that the difference between the means is not statistically significant in any of the cases ($p > 0.05$).

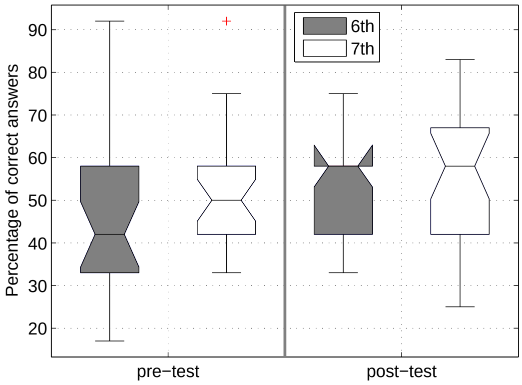


**Fig. 25** Individual students scores in the mechanical reasoning pre and post-tests

**Table 14** Post-test scores (mean and standard deviation, 6th and 7th grade comparison)

| | 6th grade (N=26) | | 7th grade (N=22) | | | |
|---|---|---|---|---|---|---|
| | M (%) | SD (%) | M (%) | SD (%) | t | p |
| pre-test | 47.15 | 11.07 | 53.50 | 14.22 | -1.2510 | 0.2173 |
| post-test | 54.50 | 18.85 | 56.50 | 14.30 | -0.5457 | 0.5879 |

### *3.2.3 Level differences*

Fig. 25 aims at comparing the individual students scores obtained in both grade levels. Notice that the difference between 6th and 7th grade is clear in the pre-test (the 7th students scores are higher than the 6th students scores), but both grade levels obtain similar scores in the post-test.

Table 14 contains the mean and standard deviation of the scores obtained by the 6th and 7th grade students in both tests. Obtained results coincide with the ones observed in Fig. 25. The 7th grade students obtain a higher mean than the 6th grade students in the pre-test. However, both means are similar in the post-test. The $p$-values given by the $t$-test allow to conclude that the difference between the mean obtained in each grade level is not statistically significant ($p = 0.2173$ and $p = 0.5879$, in the pre- and post-tests, respectively).

## 3.3 Discussion

The reported results show that the spatial ability of the participants is definitely improved. The mean scores obtained in the all the post-tests are always higher than the ones obtained in the pre-tests. Specifically, in the 6th grade case, the mean difference is statistically significant in the sub-test 3 and considering the average of all the sub-tests. In the 7th grade case, the mean difference is statistically significant in the sub-test 3 and nearly statistically significant in the sub-test 1 and on average ($p = 0.0586$ and 0.0753, respectively). Notice that in the pilot project, the mean difference was statistically significant in the sub-test 2 and 3 and nearly statistically significant on average ($p = 0.07$). Unlike in the current research, in the pilot project, the scores corresponding to the post-test of sub-test 4 (*Cube Comparisons Test*) were lower than the ones of the pre-test.

Regarding the gender differences, results reveal that, on average, the male participants outperform the female participants. However, the obtained scores depend on the sub-test. Recall that each of the sub-tests corresponds to a factor of spatial ability. Sub-test 1 and 4, which corresponds to spatial relations factor, are performed certainly better by the male participants. Notice that mental rotations are needed in order to perform these sub-tests. These results coincide with the conclusions exposed in [18], where they affirmed that the sex differences are evident for the mental rotation. Indeed, there exists a relatively large variation in the effect of sizes of gender differences in mental rotation ability across studies [23]. In addition, both [18] and [23] commented that their results indicates that the performance gap on the mental rotation ability between male and female participants tend to be small when they have more time assigned to solve and item with time limits. Therefore, they concluded that the gender difference was largest when a shorter time limit was set per item. Remark that in the current research the participants have a time limit per sub-test.

In the sub-test 2, the differences between sexes are not as clear. Actually, female participants obtain sometimes better scores. Again, this result coincides with the conclusions presented in [18]. They show that there are no differences in the *Paper Folding Test*, which they used to evaluate the spatial visualization factor, as in the current research.

With respect to mechanical reasoning, it should be noticed that the scores obtained in the pre- and post-tests are in general low. Notwithstanding, the obtained results evidence an improvement on the mechanical reasoning of the participants. However, this improvement is not statistically significant in any of the cases. One of the possible reasons is the difficulty of the test used to evaluate the mechanical reasoning of the students.

Scores obtained by the 7th grade participants are in general higher than the ones obtained by the 6th grade participants. This difference is clearer in the spatial ability performance (it is statistically significant in sub-tests 1 and 3 and on average). In the case of mechanical reasoning, the difference between 6th and 7th grades is not as notable.

As mentioned above, none of the activities included in the tests used in the spatial ability and mechanical reasoning evaluations are directly addressed in the STEM course.

Another handicap this research has to deal with is the small number of participants. It would be easier to obtain statistically significant differences with a higher number of participants. In the case of gender differences study, this disadvantage is even more critical, since the number of participants is approximately halved.

## 4 Conclusions

This paper aims at studying if the spatial ability and the mechanical reasoning of the students that attended a STEM course were improved. First of all, a hands-on learning based course labeled as STEM was designed. It was implemented for the first time at school in a 6th grade class (primary school) and in a 7th grade class (secondary school).

In order to evaluate the spatial ability of the students, a pre-test and a post-test were prepared. These tests were composed of 4 sub-tests, taking into account the most suitable and effective tests existing in the research literature. Analogously, a pre-test and a post-test were prepared to evaluate the mechanical reasoning of the students. The current research provides an extensive analysis of the obtained results.

Reported results showed that the overall performance of the students in spatial ability depends on the specific sub-test, but on average, their spatial ability was unquestionably improved. Indeed, this improvement was statistically significant. Results also manifested the improvement of the mechanical reasoning of the participants of the STEM course. However, in this case, the improvement was not statistically significant.

With respect to gender differences, obtained results evidenced that, in general, the performance does not depend on the sex of the participants. Finally, results revealed the statistically significant difference of spatial ability between 6th grade and 7th grade students, being greater in latter than in the former case. This difference was not as evident in the case of mechanical reasoning, in which case the difference between levels was not statistically significant.

As a future work, we would like to consider a larger sample size. It could be easier to obtain more statistically significant results. This fact is decisive in the case of gender differences study, when the sample is approximately halved. Furthermore, it would be interesting to find another test to measure the mechanical reasoning of the students. Scores obtained in both grades suggest that perhaps it was too difficult for the students.

## References

1. Principles and Standards for School Mathematics. National Council of Teachers of Mathematics (NCTM) (2000)
2. National Governors Association. Building a science, technology, engineering and math agenda. Retrieved from: `https://www.nga.org/files/live/sites/NGA/files/pdf/2007/0702INNOVATIONSTEM.PDF` (2007). Accessed: August 2017
3. Andersen, L.: Visual-spatial ability: Important in STEM, ignored in gifted education. Roeper Review **36**(2), 114–121 (2014)
4. Bakker, M.: Spatial Ability in Primary School: Effects of the Tridio Learning Material, Master Thesis of Psychology. University of Twente, Enschede (2008)

5. Bell, D.: The reality of STEM education, design and technology teachers' perceptions: a phenomenographic study. International Journal of Technology and Design Education **26**, 61–79 (2016)
6. Bennet, G.: Bennett Mechanical Comprehension Test. Pearson, San Antonio, USA (2008)
7. Blind:
8. Board, N.S.: Preparing the next generation of STEM innovators: identifying and developing our nation's human capital. National Science Foundation, Washington (2010)
9. Bybee, R.: The case for STEM education: Challenges and opportunities. NSTA Press, Arlington (2013)
10. Carrol, J.: Human cognitive abilities: A survey of factor-analytic studies. Standford University, School of Education Aptitude Research Project, Standford, CA (1993)
11. Contero, M., Naya, F., Company, P., Saorn, J.L.: Learning support tools for developing spatial abilities in engineering design. International Journal of Engineering Education **22**(3), 470–477 (2006)
12. Cuperman, D., Verner, I.: Fostering analogical reasoning and design skills through creating bio-inspired robotic models. In: Procedia Conference on Life Cycle Engineering 36 (CIRP), pp. 285–290 (2015)
13. Ekstrom, R.B., French, J.W., Harman, H.H., Dermen, D.: Kit of factor referenced cognitive tests. Educational Testing Service, Princeton, New Jersey (1976)
14. French, J.: The description of aptitude and achievement tests in terms of rotated factors. University of Chicago Press, Chicago, USA (1951)
15. Hegarty, M., Waller, D.: A dissociation between mental rotation and perspective-taking spatial abilities. Intelligence **32**, 175–191 (2004)
16. Humphreys, L., Lubinski, D., Yao, G.: Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. Journal of Applied Psychology **78**(2), 250–261 (1993)
17. Lin, H.: Influence of design training and spatial solution strategies on spatial ability performance. International Journal of Technology and Design Education **26**, 123–131 (2016)
18. Linn, M., Petersen, A.C.: Emergence and characterization of sex differences in spatial ability: A meta-analysis. Child Development **56**(6), 1479–1498 (1985)
19. Liu, A., Schunn, C.: Applying math onto mechanisms: mechanistic knowledge is associated with the use of formal mathematical strategies. Cognitive Research: Principles and Implications **2**(6) (2017)
20. Lohman, D.: Spatial ability: A review and reanalysis of the correlational literature. Cambridge University Press, New York, USA (1979)
21. Lord, T.: Spatial teaching. The Science Teacher **54**(2), 21–34 (1987)
22. Lubinski, D.: Spatial ability and STEM: A sleeping giant for talent identification and development. Personality and Individual Differences **459**, 344–351 (2010)
23. Maeda, Y., Yoon, S.: A meta-analysis on gender differences in mental rotation ability measured by the purdue spatial visualization tests: Visualization of rotations (PSVT:R). Educational Psychology Review **25**(1), 69–94 (2013)
24. McGee, M.G.: Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. Psychological Bulletin **86**(5), 889–918 (1979)
25. McKenna, A., Agogino, A.: Supporting mechanical reasoning with a representationally-rich learning environment. Journal of Engineering Education pp. 97–104 (2004)
26. McMunn, R.: Mechanical Comprehension Test. Sample mechanical comprehension test questions and answers. How2become Ltd, Great Britain (2012)
27. Metz, S., Donohue, S., Moore, C.: Spatial skills: a focus on gender and engineering. Resource document. `http://www.engageengineering.org/associations/11559/files/ARP_SpatialSkills.pdf` (2012). Accessed: June 2014
28. Ritz, J.M., Fan, S.C.: STEM and technology education: international stat-of-the-art. International Journal of Technology and Design Education **25**, 429–451 (2015)
29. Sorby, S.A.: Educational research in developing 3-d spatial skills for engineering students. International Journal of Science Education **31**(3), 459480 (2009)
30. Sutton, K., Williams, A.: Spatial cognition and its implications for design. International Association of Societies of Design Research, Hong Kong, China (2007)

31. Verner, I.: Robot manipulations: a synergy of visualization, computation and action for spatial instruction. International Journal of Computers for Mathematical Learning **9**, 213–234 (2004)